Mathematical Foundations for Reinforcement Learning

Spring 2023, STAT 9910-302

1 Basic course information

Instructor:	Yuting Wei; `https://yutingwei.github.io/` `ytwei@wharton.upenn.edu`
Lecture:	MW 1:15-3:45pm; Jan 11 (first lecture) to Apr 26
Location	TBD
Office hours:	Yuting Wei, by appointment, location: ARB 307

Synopsis: Reinforcement learning (RL), which is frequently modeled as sequential decision making in the face of uncertainty, has garnered growing interest in recent years due to its remarkable success in practice. Despite decades-long research efforts, however, the statistical underpinnings of RL remain far from mature, especially when it comes to sample-starved large-dimensional regimes that are of crucial operational value in practice. An explosion of research has been conducted over the past few years towards advancing the statistical frontiers of these topics, which leverage toolkits that lie at the heart of statistics, such as high-dimensional statistics, stochastic approximation, uncertainty quantification, exploration-exploitation trade-offs, and statistical learning theory.

This Ph.D. topic course aims to present a coherent framework that covers important algorithmic developments in modern RL, highlighting the connections between new ideas and classical topics. Employing Markov Decision Processes (MDPs) as the central mathematical framework, we will cover multiple important scenarios including but not limited to the simulator setting, online RL, offline RL, and multi-agent RL, gravitating our discussions around issues such as sample complexity, computational efficiency, function approximation, as well as information-theoretic and algorithmic-dependent lower bounds. Our main goal is to provide Ph.D. students with necessary background to understand the ongoing literature on RL and provide motivated students with the core toolkits for working on related research.

Prerequisites: This course is intended for Ph.D. students with strong mathematical background. There are no formal prerequisites, but students are expected to be comfortable with linear algebra, basic probability tools such as concentration inequalities, and have basic knowledge of convex optimization, such as convex functions and gradient methods.

2 Course descriptions

Tentative list of topics, some might be assigned as reading materials:

- Introduction & background
  - Markov decision processes (MDPs)
  - Classical dynamic programming
- RL under a generative model

- Model-based (plug-in) approach
- Model-free approach (e.g. stochastic approximation, Q-learning)
- Information theoretic lower bounds

- Online RL
  - Optimism principle (e.g., upper confidence bound (UCB))
  - Model-based approach (e.g., UCB value iteration)
  - Model-free approach (e.g., Q-learning with UCB)
  - Minimax regret lower bounds
  - Linear function approximation
- Offline RL
  - Pessimism principle (e.g., lower confidence bound (LCB))
  - Model-based offline RL with LCB
  - Model-free offline RL with LCB
- Policy optimization
  - Policy gradient methods
  - Natural policy gradient methods
- Multi-agent RL
  - Markov games
  - Curse of multi-agents and adversarial learning
  - Minimax lower bounds

Textbook: There is no textbook for this class. Here are some references and additional papers will be uploaded as the class proceeds.

- *Reinforcement Learning: Theory and Algorithms (draft)*, by Alekh Agarwal, Nan Jiang, Sham M. Kakade, Wen Sun
- *Reinforcement learning: An introduction*, by Richard S. Sutton, Andrew G. Barto
- *Reinforcement learning and optimal control*, by Dimitri P. Bertsekas
- *Bandit Algorithms*, by Tor Lattimore, Csaba Szepesvari

Mathematical tools:

- *High-dimensional statistics: A non-asymptotic viewpoint*, by Martin Wainwright
- *High dimensional probability: An introduction with applications in Data Science*, by Roman Vershynin
- *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, by Aharon Ben-Tal and Arkadi Nemirovski.

3 Graded components

There will be optional homework assignments, which will be posted on Canvas. Your grade for this course will be mainly determined by a final project and presentation. The final project can be either a literature review or include original research.

- Literature review. We will provide a list of related papers that are not covered in the lectures, and the literature review should involve in-depth summaries and exposition of one of these papers.
- Original research. It can be either theoretical or experimental, with the approval from the instructor. If you choose this option, you can do it either individually or in groups of two.

Three timestamps for the course project:

- Proposal (tentative: Mar 1). Submit a short report stating the papers you plan to survey or the research problems that you plan to work on. Describe why they are important or interesting, and provide some appropriate references. If you elect to do original research, you are en- couraged to connect this project with your current research (but is still related to our course content). Please do not propose an overly ambitious project. You will receive feedback from the instructor.
- In-class presentation (last week of the semester).
- A written report (tentative: May 2). You are expected to submit a final project report summarizing your findings / contributions.